

项目介绍：

项目想要研究一种能够利用 POI 数据或者路网数据对人的基站轨迹数据（低采样率）进行语义赋予，利用文本索引使得能够更加快速的处理轨迹数据，不用考虑路网的拓扑结构，同时能够利用文本的相似度或文本的分析方式来分析轨迹，对轨迹的展示更加详细。利用集成式的界面不仅能够提供快速的轨迹查询，也可以提供展示轨迹的不确定性及其原因。帮助用户分析。

关键的问题如下

- 语义赋予：

- 街道还是 POI：

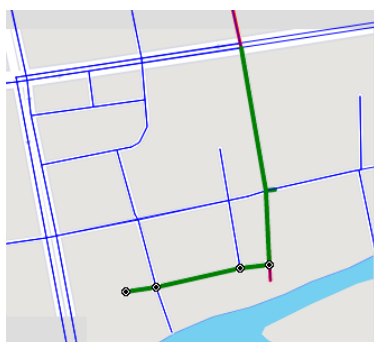
街道赋予轨迹语义：比较稀疏，可能不准，因为是范围的 voronoi 图，用街道赋予轨迹语义可能会出现偏差。



人的轨迹和路网如上图所示，人的轨迹是基于基站的位置轨迹移动图。路网数据有的地方比较密集，有的地方比较稀疏。用街道来赋予轨迹语义的难点在于：

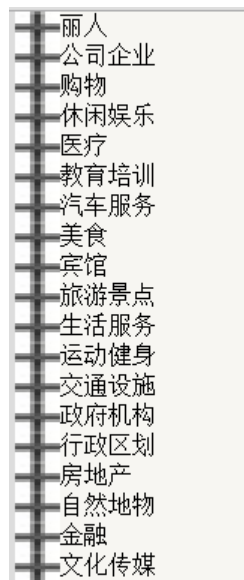
1. 一个区域内可能有很多路，到底赋予那一条道路。
2. 还有的区域可能没有道路，例如上图下方区域。
3. 道路的语义信息可能不太好帮助用户理解轨迹信息，相比于路网，如果赋予的文本信息是 POI 数据，那么用户可能更好理解轨迹的含义。

但是虽然路网不利于理解轨迹，但是在轨迹还原这一点上还是有很大优势的，而且由于道路数目相对较少，赋予语义后查询速度可能也比较快。我初步做了一点轨迹还原的工作，用户在地图上任意选择两个点，系统可以还原出这两点之间的最短路网路径和他们周围的路网。（背景地图）



这部分可以作为可视化方面的工作，找到轨迹可能的路网轨迹，现在找到的是最短轨迹，我们可以根据轨迹的匹配程度，把可能的轨迹找出来可视化。用透明度或粗细的方式来表示不同可能性大小的轨迹。

POI 赋予轨迹语义：POI 数据量比较密集，城市的每一个角落都有很多。而且 POI 是分成类别的，大类如下（每个大类都有很多小类）：



用 POI 来赋予语义轨迹的难点如下：

1. 每一个区域都有大量的 POI，在赋予轨迹的时候难以确定到底是哪个 POI 决定这个轨迹。
 - a) 一种做法是用时间来约束，比如中午晚上就是商场餐厅来赋予，白天就是上午来赋予。
 - b) 还有做法就是用户选择感兴趣的 POI 类型，用该 POI 类型来赋予语义轨迹。
 - c) 还有做法是利用该区域的 POI 统计信息来赋予，这里分别赋予局部的统计信息(eg: 该区域有多少餐馆)和全局统计信息(该区域餐馆占全部的百分比)，这样当用户查询下午 2 点去吃饭的人可能有很好的查询效果，因为我们相当于给区域赋予了职能。
 - d) 另外就是轨迹的停留点可以赋予比较多的语义信息而行进中的点可以赋予比较关键的建筑物。因为我们关注路径大都习惯看他经过了那些标志建筑，而 OD 时才会关注比较详细的 POI
2. 查询的可能性问题，由于大量的 POI 数据无法全部赋予语义，或者我们无法确定轨迹到底经过了那些 POI，查询出的结果会出现不确定性的问题，我们需要量化和可视化这些不确定性。最大子序列和地理位置的准确性可能可以来量化这些不确定性。（想用文本的方式来量化）

■ 可视化的贡献点：

系统能够赋予语义轨迹然后用文本方式来处理，那么可视化方面的贡献点有哪些。

1. 首先是不确定性的可视化，上面提到的轨迹在路网上的轨迹是不确定的，OD 的到达 POI 也是不确定的，我们可以进行量化并可视化。
2. 同时在分析轨迹的过程中，我们想要理解轨迹，以往的方式都是呈现轨迹，由于我们赋予了轨迹语义，我们可以对轨迹的语义进行可视化。轨迹经过了那些街道 POI（不确定性的），他的目的地区域职能有哪些。他的活动规律有些（白班，白夜班，自由职业，旅游）
3. 查询方面，需要设立一种比较好的对于语义的查询方法。最重要的是方便快捷的定义条件进行查询